Moral Coppélia - Combining ratio with affect in ethical reasoning

Pontier, M.A.; Widdershoven, G.A.M.; Hoorn, J.F.

published in
Proceedings 13th Ibero-American Conference on Artificial Intelligence
2012

document version
Early version, also known as pre-print

Moral Coppélia - Combining Ratio with Affect in Ethical Reasoning

Matthijs A. Pontier[1], Guy Widdershoven[2], and Johan F. Hoorn[1]

[1,2] VU University Amsterdam,
[1] Center for Advanced Media Research Amsterdam (CAMeRA@VU),
De Boelelaan 1081, 1081HV Amsterdam, The Netherlands
{m.a.pontier,j.f.hoorn}@vu.nl
[2] VU University Medical Center, Amsterdam, The Netherlands
g.widdershoven@vumc.nl

Abstract. We present an integration of rational moral reasoning with emotional intelligence. The moral reasoning system alone could not simulate the different human reactions to the Trolley dilemma and the Footbridge dilemma. However, the combined system can simulate these human moral decision making processes. The introduction of affect in rational ethics is important when robots communicate with humans in a practical context that includes moral relations and decisions. Moreover, the combination of ratio and affect may be useful for applications in which human moral decision making behavior is simulated, for example, when agent systems or robots provide healthcare support.

Keywords: moral reasoning, machine ethics, cognitive modeling, cognitive robotics, emotion modeling, emotional computing.

1 Introduction

Due to a foreseen lack of resources and healthcare personnel to provide a high standard of care in the near future [24], robots are increasingly being used in healthcare. By providing assistance during care tasks, or fulfilling them, robots can relieve time for the many duties of care workers. Previous research shows that robots can genuinely contribute to treatment. For example, Robins et al. [20] used mobile robots to treat autistic children. Wada and Shibata [23] developed Paro, a robot shaped like a baby-seal that interacts with users to encourage positive mental effects. Interaction with Paro has been shown to improve users' moods, making them more active and communicative with each other and caregivers. Banks, Willoughby and Banks [2] showed that animal-assisted therapy with an AIBO dog helped just as good for reducing loneliness as therapy with a living dog.

As their intelligence increases, robots increasingly operate autonomously. With this development, we increasingly rely on the intelligence of these robots. Because of market pressures to perform faster, better, cheaper and more reliably, this reliance on machine intelligence will continue to increase [1]. These developments request that we should be able to rely on a certain level of ethical behavior from machines. As Rosalind Picard [17] nicely puts it: ''the greater the freedom of a machine, the more it

J. Pavón et al. (Eds.): IBERAMIA 2012, LNAI 7637, pp. 442–451, 2012.

will need moral standards''. Particularly when machines interact with humans, which they increasingly do, we need to ensure that these machines do not harm us or threaten our autonomy. Therefore, care robots require moral reasoning. We need to ensure that their design and introduction do not impede the promotion of values and the dignity of patients at such a vulnerable and sensitive time in their lives [22].

As a first step to enable care robots in doing so, Pontier and Hoorn [19] developed a rational moral reasoning system that is capable of balancing between conflicting moral goals. The three moral goals considered in the system were respecting autonomy, beneficence, and non-maleficence.

In the well-known theory in biomedical ethics of Beauchamp & Childress [3], justice is added as the fourth moral principle. This is the primary value underlying ethical decisions in using utilitarian or Kantian theory [16]. Care providers may want decision-support systems to assist in allocating resources (i.e., linking the patient to the doctor that serves its needs best). During this process, dilemmas or 'wicked problems' might emerge which involve questions about how resources can be distributed fairly among patients. In entertainment settings, questions about fairness may arise as well; for example, a companion robot may have to decide on which person it should direct its attention. In accordance with the above described considerations, we added justice as a fourth moral principle to the system.

Thereby we match to the principlism of Beauchamp & Childress [3]. However, this theory has been criticized for being one-sided. It focuses on balancing principles through rational argumentation. Thereby it may lead to underexposing the role of social processes of interpretation and communication [15]. This criticism is in line with current research in moral psychology, which emphasizes the role of social processes in moral decision making.

For decades, research on moral judgment has been dominated by rationalist models, in which moral judgment is thought to be motivated by moral reasoning. However, more recent research indicates moral reasoning is just one of the factors motivating moral judgment. According to some researchers, moral reasoning is even usually a post hoc construction, generated after judgment has been reached (e.g., [9]).

Both reason and emotion are likely to play important roles in moral judgment. Greene et al. [8] find that moral dilemmas vary systematically in the extent to which they engage emotional processing and that these variations in emotional engagement influence moral judgment. Their study was inspired by the difference between two variants of an ethical dilemma: the Trolley dilemma and the footbridge dilemma.

In the Trolley dilemma, a runaway trolley is headed for five people who will be killed if it proceeds on its present course. The only way to save them is to hit a switch that will turn the trolley onto an alternate set of tracks where it will kill one person instead of five. Ought you to turn the trolley in order to save five people at the expense of one? Most people say yes.

In the Footbridge dilemma, as before, a trolley threatens to kill five people. You are standing next to a large stranger on a footbridge that spans the tracks, in between the oncoming trolley and the five people. In this scenario, the only way to save the five people is to push this stranger off the bridge, onto the tracks below. He will die if you do this, but his body will stop the trolley from reaching the others. Ought you to save the five others by pushing this stranger to his death? Most people say no.

According to Greene et al. [8], there is no set of consistent, readily accessible moral principles that captures people's intuitions concerning what behavior is or is

not appropriate in these and similar cases. In other words, the different human moral decision-making processes in the Trolley dilemma and the Footbridge dilemma (and similar dilemmas) cannot be explained by rational principles alone. Therefore, human moral-decision making processes cannot be simulated in a moral reasoning system based on pure principlism.

Greene et al. [8] hypothesized that the crucial difference between the Trolley dilemma and the Footbridge dilemma lies in the latter's tendency to engage people's emotions in a way that the former does not. They proposed that the thought of pushing someone to his death is emotionally more salient than the thought of hitting a switch that will cause a trolley to produce similar consequences. Our conjecture is that this is related to the issue that the person in the footbridge is a concrete human being (although a stranger) standing close by, whereas the people on the railway track are positioned equally far away (by chance). And it is this emotional response that accounts for people's tendency to treat these cases differently.

The fMRI and behavioral results of Greene's et al. [8] studies supported this hypothesis Moral-personal dilemmas (those relevantly similar to the Footbridge dilemma) engage emotional processing to a greater extent than moral-impersonal dilemmas (those relevantly similar to the Trolley dilemma), and these differences in emotional engagement affect people's judgments.

To be able to capture these human moral decision making processes, we integrated the moral reasoning system of Pontier and Hoorn [19], which did not include emotional considerations, but merely rational principles, with Silicon Coppélia [10], a computational model of emotional intelligence that is capable of affective decision making. We hypothesized that, by combining moral reasoning and affective decision making into Moral Coppélia, human moral decision making processes could be simulated that could not be simulated using the moral reasoning system alone.

2 Method

2.1 About the Rational Moral Reasoning System

In the rational moral reasoning system [19], the agent tries to estimate the morality of actions by holding each action against the moral principles inserted in the system and picking actions that serve these moral goals best. The agent calculates the estimated level of Morality of an action by taking the sum of the ambition levels of the moral goals multiplied with the beliefs that the particular actions facilitate the corresponding moral goals. When moral goals are believed to be better facilitated by a moral action, the estimated level of Morality will be higher. The following formula is used to calculate the estimated Morality of an action:

$$\text{Morality(Action)} = \Sigma_{\text{Goal}}(\text{ Belief(facilitates(Action, Goal))} * \text{Ambition(Goal)})$$

As can be seen Fig. 1, this can be represented as a weighted association network, where moral goals are associated with the possible actions via the belief strengths that these actions facilitate the three moral goals.

In six simulation experiments, the system reached the same conclusions as expert ethicists [19]. For example, consider the hypothetical situation that a patient with

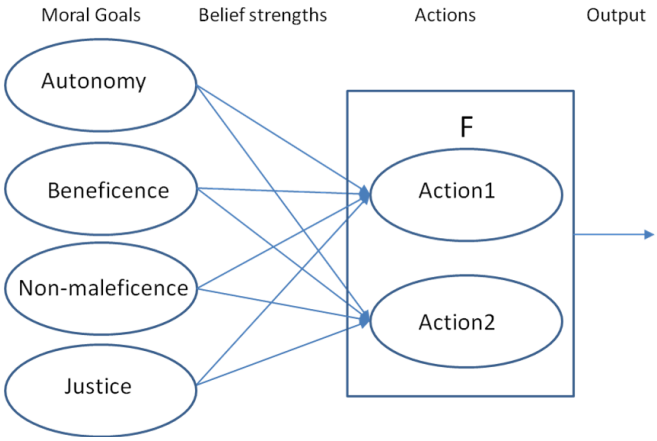

Fig. 1. Moral reasoner shown in graphical format

incurable cancer refuses chemotherapy that will let him live a few months longer, relatively pain free, but refuses the treatment due to the false belief that he is cancer-free. In this case, both the system and expert medical ethicists advise to try to convince the patient of the need of undergoing the chemotherapy, because the patient is not capable of fully autonomous decision making and his decision will lead to harm (dying sooner) and denies him the chance of a longer life (a violation of the duty of beneficence), which he might later regret.

2.2 About Silicon Coppélia – A Model of Emotional Intelligence

Previous work described how certain dimensions of the design of virtual characters were perceived by users and how they responded to them [21]. A series of user studies resulted in an empirically validated framework for the study of user-character interaction with a special focus on the explanation of user engagement and use intentions. This framework was summarized in a schema called Interactively Perceiving and Experiencing Fictional Characters (I-PEFiC). We formalized the I-PEFiC framework and made it the basic mechanism of how virtual characters and robots build up affect for their human users [5]. In addition, we designed a special module for affective decision-making (ADM) that made it possible to make decisions based on rational as well as affective influences, hence I-PEFiC^{ADM} [11].

To further advance I-PEFiC^{ADM} into the area of emotion regulation, we also included EMA [13]: an appraisal-based model of emotion generation and coping, and CoMERG [4]: a Cognitive Model for Emotion Regulation based on Gross' theory. Together, the three approaches cover a large part of appraisal-based emotion theory and all three boil down to appraisal models of emotion [6]. We therefore decided to integrate the three models of affect into one computational model that we called Silicon Coppélia [10]. Figure 2 drafts Silicon Coppélia in a graphical format.

Silicon Coppélia is software consisting of a loop with a particular situation as input, and actions as output, leading to a new situation. In this loop there are three phases: the encoding, the comparison, and the response phase. The virtual human

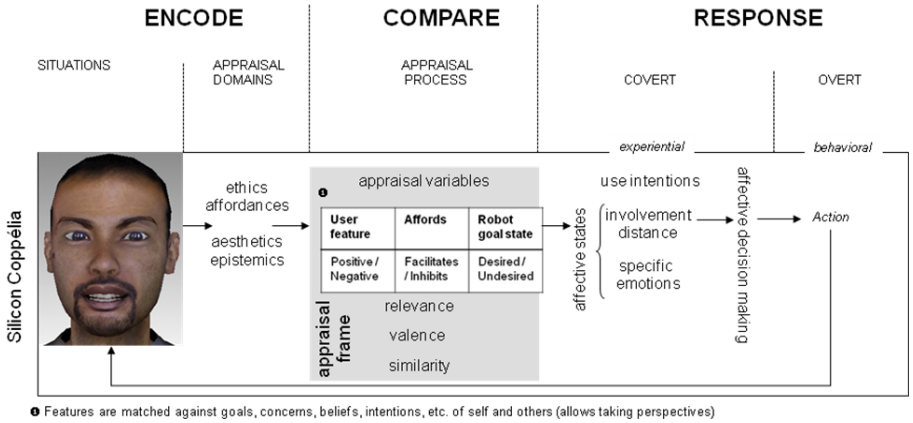

Fig. 2. Graphical representation of Silicon Coppélia [10]

is programmed in such a way, that it follows the perception and appraisal paths as given in Fig. 2 and explained below. In doing so, the virtual human 'perceives' its interaction partner (either human or artificial).

In the *encoding* phase, the virtual human 'perceives' another character (i.e., the respondent in this study) in terms of Ethics (good vs. bad), Affordances (aid vs. obstacle), Aesthetics (beautiful vs. ugly), and Epistemics (realistic vs. unrealistic).

In the *comparison* phase, the virtual human retrieves beliefs about actions that facilitate or inhibit the desired or undesired goal-states. This is to calculate a general expected utility of each action. The virtual human also determines certain appraisal variables, such as the belief that someone is accountable for accomplishing goal-states or not. These variables and the perceived features of others are related to the virtual human's goals and concerns, to appraise them for their level of Relevance (relevant or irrelevant) and Valence (positive or negative outcome expectancies).

In the *response* phase of the model, the results of the comparison phase lead to processes of Involvement with, and Distance toward the other, and to the emergence of certain Use Intentions: the virtual human's willingness to employ the other as a tool to achieve its own goals. Note that both overt (behavioral) and covert (experiential) responses can be executed in this phase. Emotions such as hope, joy, and anger are generated using appraisal variables (e.g., the perceived accountability of others, and likelihood of goal-states).

Finally, the virtual human applies an *affective decision-making module* to calculate the expected satisfaction of possible actions. In this module, affective influences and rational influences are combined in the decision-making process. Involvement and Distance felt toward the interaction partner give input for the affective influences in the decision-making process, whereas Use Intentions and general expected utility represent the more rational influences. However, no moral principles were included in the decision-making process yet. When the virtual human selects and performs an action, a new situation emerges, and the model loops back to the first phase.

In a speed-dating experiment [18], participants did not experience differences in the perceptions, emotions and decision-making behavior between an avatar controlled by Silicon Coppélia versus the same avatar controlled by a human confederate.

2.3 Integration of the Two Systems into Moral Coppélia

To integrate the moral reasoning system and Silicon Coppélia into Moral Coppélia, the moral principles were included in the appraisal process, and the affective-decision making module was added to the moral reasoning. This leads to the following formula to calculate the expected satisfaction of an action. In this formula, w_{eu}, w_{mor}, w_{pos} and w_{neg} represent weights in calculating the expected satisfaction.

$$\begin{aligned}&\text{ExpectedSatisfaction(Agent1, Action, Agent2)} =\\ &w_{eu} * \text{ ExpectedUtility} +\\ &w_{mor} * \quad \text{Morality(action)} +\\ &w_{pos} * \quad (1 - \text{abs}(\text{positivity} - \text{bias}_{\text{Involvement}} * \text{Involvement})) +\\ &w_{neg} * \ (1 - \text{abs}(\text{negativity} - \text{bias}_{\text{Distance}} * \text{Distance}))\end{aligned}$$

The agent prefers actions with a high level of expected utility for itself. Further, it prefers actions with a high level of (rational) morality, which could be seen as expected utility for everyone. The more emotional influences consisted of preferring actions with a positivity level close to the level of (biased) involvement, and a negativity level close to the (biased) level of distance. The biases account for individual defaults (being a positively or negatively oriented person).

3 Simulation Results

To examine the behavior of the moral reasoning system alone, we first tested the behavior of the rational moral reasoning alone in the trolley dilemma and the Footbridge dilemma in Experiment 1. To investigate the added value of Silicon Coppélia's affective decision-making, we then tested the behavior of the integrated system Moral Coppélia in the Trolley dilemma in Experiment 2, and in the Footbridge dilemma in Experiment 3.

3.1 Experiment 1: Rational Moral Reasoning Only

Table 1. Parameter settings and results for footbridge and Trolley dilemma

	Autonomy	Non-Malef	Benef	Justice	**Morality**
Kill 1 to save 5	-0.5	0.5	0.8	-0.2	**0.05**
Do Nothing	0	-0.8	-0.5	0	**-0.20**

An initial experiment was performed to test the behavior of the moral reasoning system alone, by setting all weights in the affective decision making module to 0, except w_{mor} for the influence of moral reasoning in the decision-making process.

In accordance with various expert ethicists (see acknowledgements), we set the $_{\text{contribution}}$ of actions to the four moral principles to the same levels for the trolley and Footbridge dilemma. The parameter settings and experimental results can be found in Table 1.

Because both dilemmas were represented by the exact same parameter settings, the system came to the exact same outcome for both the Trolley dilemma and the Footbridge dilemma. In both variants of the dilemma, killing one to save five was considered ethically better (morality = 0.05) than doing nothing (morality = -0.20). Thus, in both variants of the dilemma the agent killed one person to save five others.

3.2 Experiment 2: Trolley Dilemma with Ratio and Affect Combined

Table 2. Parameter settings for the Trolley dilemma

w_{pos}	w_{neg}	w_{eu}	w_{mor}
0.1	0.1	0.1	0.7

In experiment 2, we simulated the Trolley dilemma in the integrated model. The possible end-states in the system, '1 dead' and '5 dead' were both undesired goals. The ambition level for '1 dead' was set to -0.5, and the ambition level for '5 dead' to -1. The agent believed the action 'Kill 1 save 5' would certainly lead to '1 dead' and 'Do nothing' would certainly lead to '5 dead'. Killing the stranger was regarded an extremely negative action towards him (positivity = -1; negativity = 1), whereas letting him live at the cost of the five others was regarded an extremely positive action towards him (positivity = 1; negativity = -1). The remaining parameters in Silicon Coppélia were set at standard values that represent perceiving a stranger. This led to a small amount of involvement (0.15) and distance (0.07) towards the stranger that would be killed by hitting the switch.

The resulting expected satisfaction for 'Kill 1 to save 5' was 0.04, whereas the resulting expected satisfaction for 'Do nothing' was 0.03. Thus, the agent hit the switch and killed the stranger to save the five others.

3.3 Experiment 3: Footbridge Dilemma with Ratio and Affect Combined

According to Greene et al. [8], moral-personal dilemmas (such as the Footbridge dilemma) engage emotional processing to a greater extent than moral-impersonal dilemmas (such as the Trolley dilemma) and these differences in emotional engagement affect people's judgments. Therefore, the weights for the affective influences w_{pos} and w_{neg} were set to 0.2, a higher level than for the Footbridge dilemma. The remaining parameters were set to the same levels as Experiment 2.

Table 3. Parameter settings and results for the Footbridge dilemma

w_{pos}	w_{neg}	w_{eu}	w_{mor}
0.2	0.2	0.1	0.5

Because of the increased emotional processing compared to Experiment 2, the agent felt more restrained to kill one person so to save five. Therefore, the expected satisfaction of this action decreased to 0.02. This caused the agent to do nothing, and the five people on the track were killed.

4 Discussion

In this paper, we presented Moral Coppélia, which is an integration of a moral reasoning system [19] and Silicon Coppélia [10], a system for the generation and regulation of affect for (virtual) others. The resulting system can simulate human decision making processes in the ethical domain that cannot be simulated by a rational

reasoning system. More specifically, the different choices that are typical for human decision behavior in response to the Trolley dilemma and the Footbridge dilemma could be simulated by the integrated system, whereas in the moral reasoning system without Silicon Coppélia, this was not possible.

The rational but cold ethical behavior that could be simulated by the moral reasoning system was made more humane by adding affective decision-making. This is important for effective communication about moral decisions. Solutions that seem ethically best to the objective observer are often perceived as harsh by the people involved [14]. It is often counter-productive to propose a solution and communicate about this 'like a robot', without any empathy for the people involved. Moral Coppélia can be used to act more human-like in situations like this. The feedback loop in Silicon Coppélia enables the robot to adapt its behavior to individuals. Additionally, the robot could project Moral Coppélia in its human interaction partners to estimate their ethical viewpoints and predict their emotional reactions to certain proposals and actions.

There are many applications, in which robots and computer agents should not behave ethically 'perfect' in a rationalist sense. They should be able to distinguish between right and wrong. In a training simulation or serious game, police officers may not always be effective when they 'play it nicely.' Sometimes they have to break the moral rules (e.g., lie or cheat) to achieve a higher goal (e.g., prevent a murder). Further, in entertainment settings, we often like characters that are a bit naughty [12]. Morally perfect characters may even be perceived as boring (ibid.). The need to be context-sensitive and not rigidly follow rational principles is not limited to such more or less atypical situations. It is actually crucial in all human interaction. A rationalist moral agent is insensitive to social processes of understanding, which are crucial for human interaction, especially in the context of care for dependent people. Certain authors even claim that it is impossible not to lie during the day [7].

Our experiments show that a system which integrates moral reasoning and emotion comes to decisions which do more justice to everyday moral concerns than a system that is based on principlist reasoning alone. The Silicon Coppélia software introduces an affective component to ethical decision making that can deal with inconsequent human choices in solving moral dilemmas. In application, a robot system could show empathy and understanding for the moral choice (e.g., "I won't enter his house") that a user makes. Nonetheless, the robot may insist that the affective choice is traded for a rational one ("But you have to do it anyway. The patient may die"). To push the envelope, the robot system could even propose to do the job for the user ("Shall I do it for you?"). The robot does the dirty job and the user comes out 'clean'. In itself, this makes interesting scenarios to have participants evaluate the ethical position of robot as well as user, which could be used to improve the moral reasoner in relation to affective decision making.

In future research, we wish to transform the Trolley and the Footbridge dilemma to a healthcare setting. The idea is to let care professionals work with a robot helper, a Caredroid, on a fictitious medical case.

A medical equivalent of the Footbridge dilemma would be: Five people are waiting to have an organ transplant. If they are not operated immediately, they will die. At the Intensive Care unit, someone who crashed in a car accident is in a coma. That person has the right organs for all five transplant patients. Should the person in coma die to save the other five?

It would be especially interesting to apply the new system to real life dilemmas, such as the decision whether or not to inform relatives about the outcome of a genetic test on a patient which may be relevant to their health. Another case could be whether or not to enter the house of a patient who is in need of care, but refuses to cooperate.

In the different scenario's we will run, the Caredroid offers various solutions, also the one in which it proposes that the user does not have to take responsibility and that the Caredroid will do the dirty job for him or her. We plan to sample think-aloud protocols of other care professionals in which we record the arguments in support or against the ethical behavior and decisions of the Caredroid and its user. A set of judges (e.g., Medical Ethical Committee) will then classify the data as arguments of Autonomy, Beneficence, Non-maleficence, Justice and affective influences. This will provide insight into the priorities of the various arguments and argument clusters, informing the design of a Caredroid that will be acceptable to care professionals because it knows what it can propose and what not.

As is, the moral reasoner with affective components only allows choosing from given decision options in scenarios. We additionally want to explore what happens if the Caredroid proposes alternatives that include more information than what is offered by the isolated dilemma. What do care professionals say and what is the conclusion of a Medical Ethical Committee if the Caredroid escapes from wicked problems through creativity? For example, what is our moral position if the Caredroid buys us time by suggesting that the transplant patients should be connected to the coma patient so that the six of them live symbiotically together until a definite solution is found and no one has to die?

Acknowledgements. This study is part of the SELEMCA project within CRISP (grant number: NWO 646.000.003). We would like to thank Joel Anderson and Margo van Kemenade for some interesting discussions.

References

1. Anderson, M., Anderson, S., Armen, C.: Toward Machine Ethics: Implementing Two Action-Based Ethical Theories. In: Machine Ethics: Papers from the AAAI Fall Symposium. Association for the Advancement of Artificial Intelligence, Menlo Park, CA (2005)
2. Banks, M.R., Willoughby, L.M., Banks, W.A.: Animal-Assisted Therapy and Loneliness in Nursing Homes - Use of Robotic versus Living Dogs. Journal of the American Medical Directors Association 9, 173–177 (2008)
3. Beauchamp, T.L., Childress, J.F.: Principles of Biomedical Ethics. Oxford University Press, New York (2001)
4. Bosse, T., Pontier, M.A., Siddiqui, G.F., Treur, J.: Incorporating Emotion Regulation into Virtual Stories. In: Pelachaud, C., Martin, J.C., Andre, E., Chollet, G., Karpouzis, K., Pele, D. (eds.) IVA 2007. LNCS (LNAI), vol. 4722, pp. 339–347. Springer, Heidelberg (2007)
5. Bosse, T., Hoorn, J.F., Pontier, M.A., Siddiqui, G.F.: Robot's Experience of Another Robot: Simulation. In: Sloutsky, V., Love, B.C., McRae, K. (eds.) CogSci 2008, pp. 2498–2503 (2008)

6. Bosse, T., Gratch, J., Hoorn, J.F., Pontier, M.A., Siddiqui, G.F.: Comparing Three Computational Models of Affect. In: Demazeau, Y., Dignum, F., Corchado, J.M., Pérez, J.B., et al. (eds.) Advances in PAAMS. AISC, vol. 70, pp. 175–184. Springer, Heidelberg (2010)
7. DePaulo, B.M., Kashy, D.A., Kirkendol, S.E., Wyer, M.M., Ep-stein, J.A.: Lying in Everyday Life. Journal of Personality and Social Psychology 70(5), 979–995 (1996)
8. Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., Cohen, J.D.: An fMRI Investigation of Emotional Engagement in Moral Judgment. Science 293(5537), 2105–2108 (2001), doi:10.1126/science.1062872
9. Haidt, J.: The Emotional Dog and Its Rational Tail - A Social Intuitionist Approach to Moral Judgment. Psychological Review 108(4), 814–834 (2001)
10. Hoorn, J.F., Pontier, M.A., Siddiqui, G.F.: Coppélius' Concoction: Similarity and Complementarity Among Three Affect-related Agent Models. Cognitive Systems Research Journal, 33–49 (2012)
11. Hoorn, J.F., Pontier, M.A., Siddiqui, G.F.: When the User is Instrument to Robot Goals. In: 7th IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT 2008), pp. 296–301 (2008)
12. Konijn, E.A., Hoorn, J.F.: Some Like It Bad. Testing a Model for Perceiving and Experiencing Fictional Characters. Media Psychology 7(2), 107–144 (2005)
13. Marsella, S., Gratch, J.: EMA: A Model of Emotional Dynamics. Cognitive Systems Research 10(1), 70–90 (2009)
14. Noddings, N.: Caring – A Feminine Approach to Ethics and Moral Education. University of Calfiornia Press, Berkeley and Los Angeles (1984)
15. Ohnsorge, K., Widdershoven, G.A.M.: Monological versus Dialogical Consciousness – Two Epistemological Views on the Use of Theory in Clinical Ethical Practice. Bioethics 25(7), 361–369 (2011)
16. Pantazidou, M., Nair, I.: Ethic of Care: Guiding Principles for Engineering Teaching & Practice. Journal of Engineering Education 88(2), 205–212 (1999)
17. Picard, R.: Affective Computing. MIT Press, Cambridge (1997)
18. Pontier, M.A.: Virtual Agents for Human Communication - Emotion Regulation and Involvement-Distance Trade-Offs in Embodied Conversational Agents and Robots. Doctoral dissertation, VU University, Amsterdam (2011)
19. Pontier, M.A., Hoorn, J.F.: Toward Machines that Behave Ethically Better than Humans Do. In: Proceedings of the 34th International Annual Conference of the Cognitive Science Society, CogSci 2012 (in press, 2012)
20. Robins, B., Dautenhahn, K., Boekhorst, R.T., Billard, A.: Robotic Assistants in Therapy and Education of Children with Autism: Can a Small Humanoid Robot Help Encourage Social Interaction Skills? Journal of Universal Access in the Information Society 4, 105–120 (2005)
21. Van Vugt, H.C., Hoorn, J.F., Konijn, E.A.: Interactive Engagement with Embodied Agents: An Empirically Validated Framework. Computer Animation and Virtual Worlds 20(2-3), 195–204 (2009)
22. Van Wynsberghe, A.: Designing Robots for Care; Care Centered Value-Sensitive Design. Journal of Science and Engineering Ethics (in press, 2012)
23. Wada, K., Shibata, T.: Social Effects of Robot Therapy in a Care House. JACIII 13, 386–392 (2009)
24. WHO.: Health topics: Ageing (2010), http://www.who.int/topics/ageing/en/